

Database

SCUD: *Saccharomyces Cerevisiae* Ubiquitination DatabaseWon-Chul Lee^{1,2}, Minho Lee¹, Jin Woo Jung², Kwang Pyo Kim^{*2,3} and Dongsup Kim^{*1}

Address: ¹Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea, ²Department of Molecular Biotechnology, Konkuk University, Seoul 143-701, Korea and ³Institute of Biomedical Science and Technology, Konkuk University, Seoul 143-701, Korea

Email: Won-Chul Lee - cleariv@kaist.ac.kr; Minho Lee - MinhoLee@kaist.edu; Jin Woo Jung - sildcw@konkuk.ac.kr; Kwang Pyo Kim^{*} - kpkim@konkuk.ac.kr; Dongsup Kim^{*} - kds@kaist.ac.kr

^{*} Corresponding authors

Published: 24 September 2008

Received: 26 May 2008

BMC Genomics 2008, **9**:440 doi:10.1186/1471-2164-9-440

Accepted: 24 September 2008

This article is available from: <http://www.biomedcentral.com/1471-2164/9/440>

© 2008 Lee et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Ubiquitination is an important post-translational modification involved in diverse biological processes. Therefore, genomewide representation of the ubiquitination system for a species is important.

Description: SCUD is a web-based database for the ubiquitination system in *Saccharomyces cerevisiae* (Baker's yeast). We first searched for all the known enzymes involved in the ubiquitination process in yeast, including E1, E2, E3, and deubiquitination enzymes. Then, ubiquitinated substrates were collected by literature search. Especially, E3 and deubiquitination enzymes are classified into classes and subclasses by their shared domains and unique functions. As a result, 42 different E3 enzymes were grouped into corresponding classes and subclasses, and 940 ubiquitinated substrates including mutant substrates were identified. All the enzyme and substrate information are interconnected by hyperlinks, which makes it easy to view the enzyme-specific ubiquitination information.

Conclusion: This database aims to represent a comprehensive yeast ubiquitination system, and is easily expandable with the further experimental data. We expect that this database will be useful for the research on the ubiquitination systems of other higher organisms.

SCUD is accessible at <http://scud.kaist.ac.kr>

Background

Ubiquitination is involved in many biological processes including cell signaling pathway and protein quality checkpoint system, by which ubiquitinated proteins are delivered to the 26S proteasome and degraded. Many recent studies have demonstrated that protein ubiquitination not only guarantees its degradation but also results in other consequences such as receptor endocytosis and protein transport, indicating that ubiquitination is an

important process with highly complicated biological behaviors.

Experimental results about ubiquitination of numerous substrates have been accumulated, but there is no database which can easily facilitate further experiments or computational analysis. Recently, a database, UbiProt, which contains the information about ubiquitinated substrates from diverse organisms, has been appeared [1].

However, UbiProt is highly dependent on a small number of high-throughput proteomics studies. Moreover, it fails to include many previously identified substrates. Importantly, UbiProt has little information about enzymes which are involved in the ubiquitination process.

Ubiquitination of a substrate is carried out by sequential reactions of enzymes. First, E1 ubiquitin-activating enzyme modifies a ubiquitin so that it is activated to a reactive state (making it likely that C-terminal glycine on the ubiquitin will react with the lysine side-chains on the substrate protein). Then, E2 ubiquitin-conjugating enzyme receives an activated ubiquitin, and catalyzes the attachment of the ubiquitin to the substrate protein in concert with E3 ubiquitin-protein ligase. Here, E3 enzymes are important as they determine the specificity to their substrates. Either a protein complex or a unit protein plays a role as E3 ligase. In case of E3 ligase complex, a specialized subunit protein called substrate receptor performs docking with the substrates, while the unit E3 ligase has a special domain with which the substrates interact. Therefore, it is necessary to identify all the known E3 enzymes, followed by classifying them according to their unique domain or complex type, and finally discovering E3-specific ubiquitinated substrates. If we are able to distinguish the substrates by E3 enzymes, it is possible to find sequence motifs or structural characteristics that contribute to the E3 enzyme-substrate specific interaction. Moreover, ubiquitinated site information would help identify the conditions on which specific lysines are selected and ubiquitinated. For these purposes, we searched for all the known enzymes involved in the ubiquitination process in budding yeast, *Saccharomyces cerevisiae*, and classified E3 enzymes into E3 classes and subclasses. Then, we collected ubiquitinated substrates either by E3-based literature search or by reference to high-throughput proteomics studies. This database aims to represent a comprehensive ubiquitination system in yeast.

Construction and content

We chose yeast as a model organism because a majority of known ubiquitination substrates were obtained from yeast. Moreover, the ubiquitination enzyme system is highly sophisticated and has many unrevealed aspects in other higher organisms, especially in human, making it difficult to systematically represent the enzymes involved in the ubiquitination process. Therefore, the yeast ubiquitination system appears a good platform from which the ubiquitination research can be extended to other organisms.

Identification of Ubiquitination Enzymes and E3 enzyme Classification

We searched for all known ubiquitination enzymes including E1, E2, and deubiquitination enzymes through *Saccharomyces* Genome Database (SGD, <http://www.yeastgenome.org>) and Swiss-Prot <http://www.expasy.org>. For E3 enzymes, we looked for all the proteins showing ubiquitin-protein ligase activity and extracted the proteins involved in substrate recognition independently or dependently on other proteins (in case of complexes). We called these proteins 'substrate receptors' that can act as a substrate-recognition subunit of E3 ligase complex or as E3 ligase alone. It has been known that E3 enzymes can be divided into different groups. We classified the substrate receptors according to the following E3 enzyme classes and subclasses.

- HECT (Homologous to E6-AP C terminus) class [2]
- RING (Really Interesting New Gene) class [3,4]
 - SPRF (Single Protein RING Finger) subclass: Single protein ligases containing RING domain. They possess both the substrate recognition part and the E2 interaction part on them.
 - SCF (SKP1, CDC53, and F-box) subclass: Cullin (CDC53)-dependent multi-subunit E3 enzyme complex. F-box proteins are in charge of substrate recognition.
 - ECS (ELC1, CUL3, and SOCS/BC-box) subclass: Cullin (CUL3)-dependent multi-subunit E3 enzyme complex. SOCS/BC-box proteins are involved in substrate recognition.
 - U-BOX subclass: U-BOX is similar to RING domain.
 - APC (Anaphase Promoting Complex) subclass: Multi-subunit E3 enzyme complex. Its subunit "APC" is similar to Cullins in SCF and ECS complex.

Here, U-box containing proteins were previously known as E4 enzyme which promotes the polyubiquitination [5]. However, it has been also verified that U-box proteins can intermediate the substrate ubiquitination with E1 and E2, regardless of E3 involvement. This results in defining the U-box proteins as a new family of ubiquitin ligases [4]. Diagrams of HECT class and 5 RING subclasses are described in Figure 1.

Identification of Ubiquitinated Substrates

Using each substrate receptor in each E3 enzyme class or subclass as a query, we collected by literature search all substrate information as well as E2 and deubiquitination

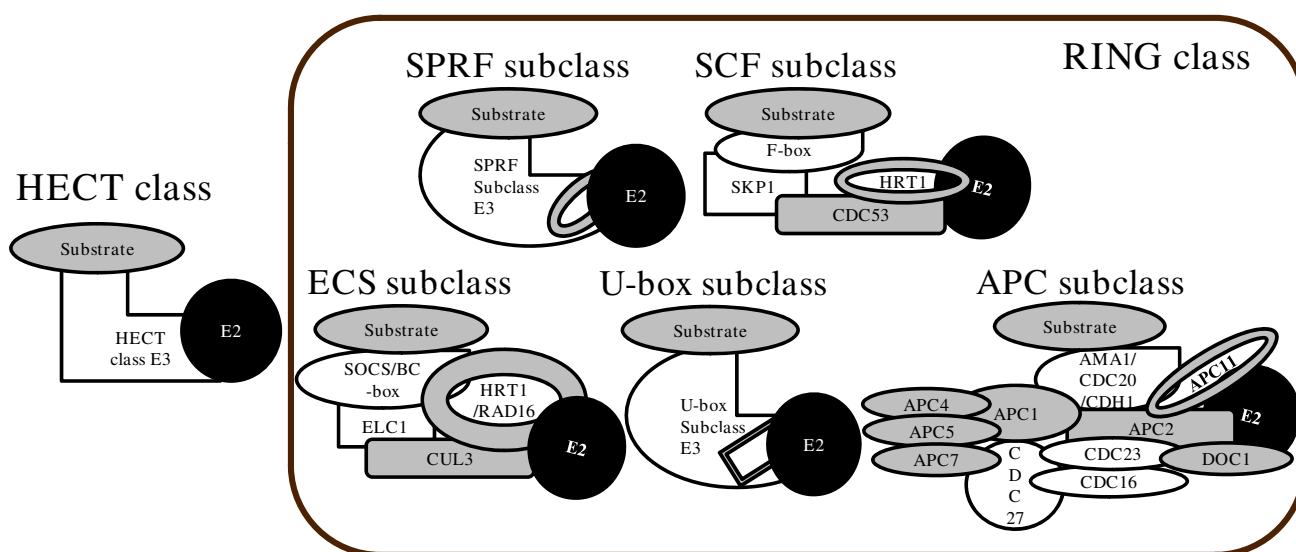


Figure 1
Diagrams of E3 enzyme classes and subclasses.

enzymes which participate in the substrate ubiquitination. The substrate information includes the ubiquitination site, type, enzymes, and ubiquitination effect as well as related references. In addition, we included the ubiquitinated substrates identified by several proteomics approaches [6-9], although most of them do not have the specific information on ubiquitination enzymes.

On the other hand, we also found the mutant and synthetic substrates which are ubiquitinated by already known E3 enzymes. In the endoplasmic reticulum (ER), two representative E3 enzymes (HRD1 and SSM4) are responsible for the ubiquitination of most ER substrates [10]. HRD1 and SSM4 target the misfolded mutant ER proteins as well as wild-type proteins. Moreover, UFD (Ubiquitin Fusion Degradation) pathway protein UFD2 and UFD4 are involved in the ubiquitination of synthetic ubiquitin fusion proteins such as Ub^{V76}-V-βgal, while their physiological substrates are mostly unknown [11].

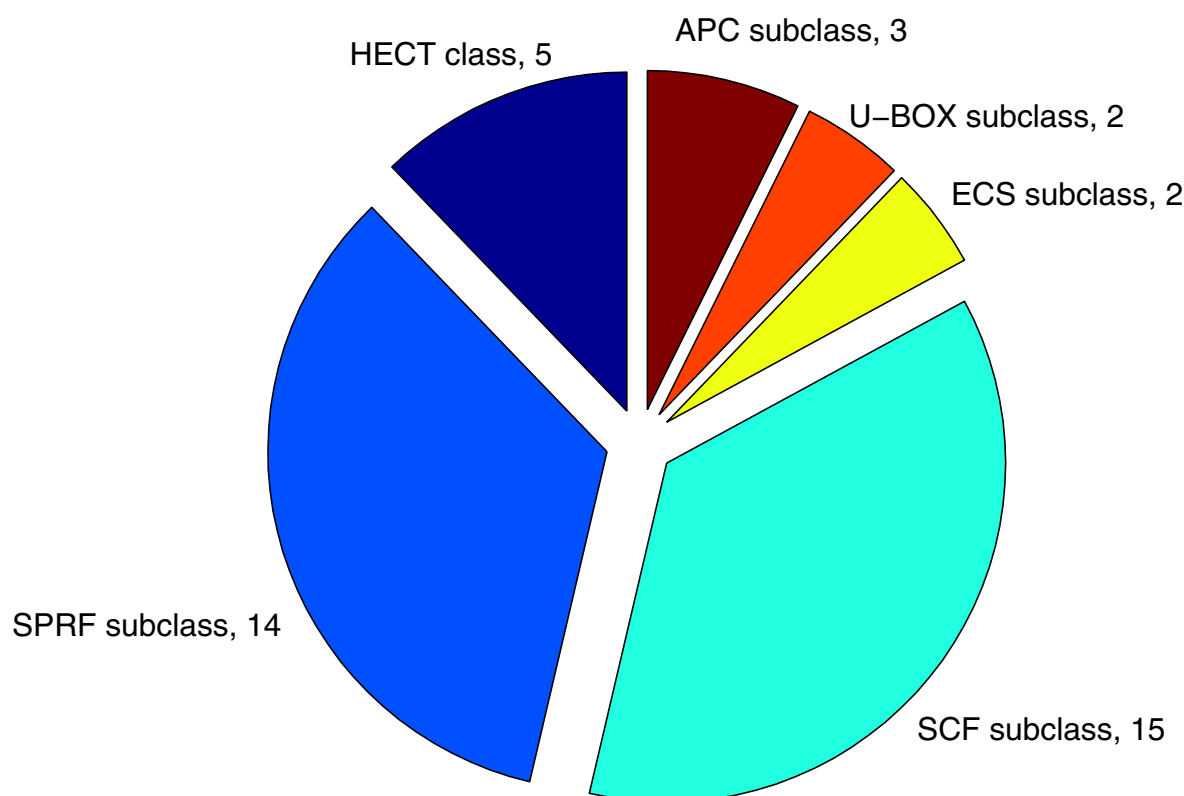
Database Content

We found total 1 E1, 11 E2, 42 E3, 20 deubiquitination enzymes, and 940 ubiquitinated substrates including mutant and synthetic substrates (see Additional file 1). The E3 enzymes are further grouped into different E3 classes and subclasses (Figure 2). Most E3 enzymes possess their own substrates, while the substrates of some E3 enzymes are not known yet.

Utility and discussion

The interfaces of our database were implemented using MySQL, PHP, AJAX, and Python. The database displays a list of all known E1, E2, E3, and deubiquitination enzymes found in the yeast ubiquitination system. Particularly, E3 and deubiquitination enzymes are grouped based on their shared domains or distinct functions. The detailed information on an individual enzyme is provided by a hypertext link. It includes general protein information as well as a list of substrates ubiquitinated by the enzyme (Figure 3). Most E3 enzymes possess their own substrates, while the substrates of some E3 enzymes are not known yet. Similarly, a list of all known substrates and individual substrate information are also available. The individual substrate information involves ubiquitination site, type, enzymes, receptors, effect, and general protein information (Figure 3). All the enzymes and substrate information are associated by hyperlinks.

On the other hand, it is necessary to provide the experimental evidences by which proteins are regarded as ubiquitin-conjugates to ensure the quality of identified substrates. For example, it is well known that ubiquitin-modified proteins are likely to be degraded by proteasome or vacuole. Therefore, sometimes proteins showing stability change are assumed to be ubiquitinated. However, the stability change can be explained by other reasons except for the ubiquitination. To offer the evidences of ubiquitination, we collected all the experimental evidences of ubiquitinated proteins in the references, categorized them, and included in the database along with

**Figure 2**

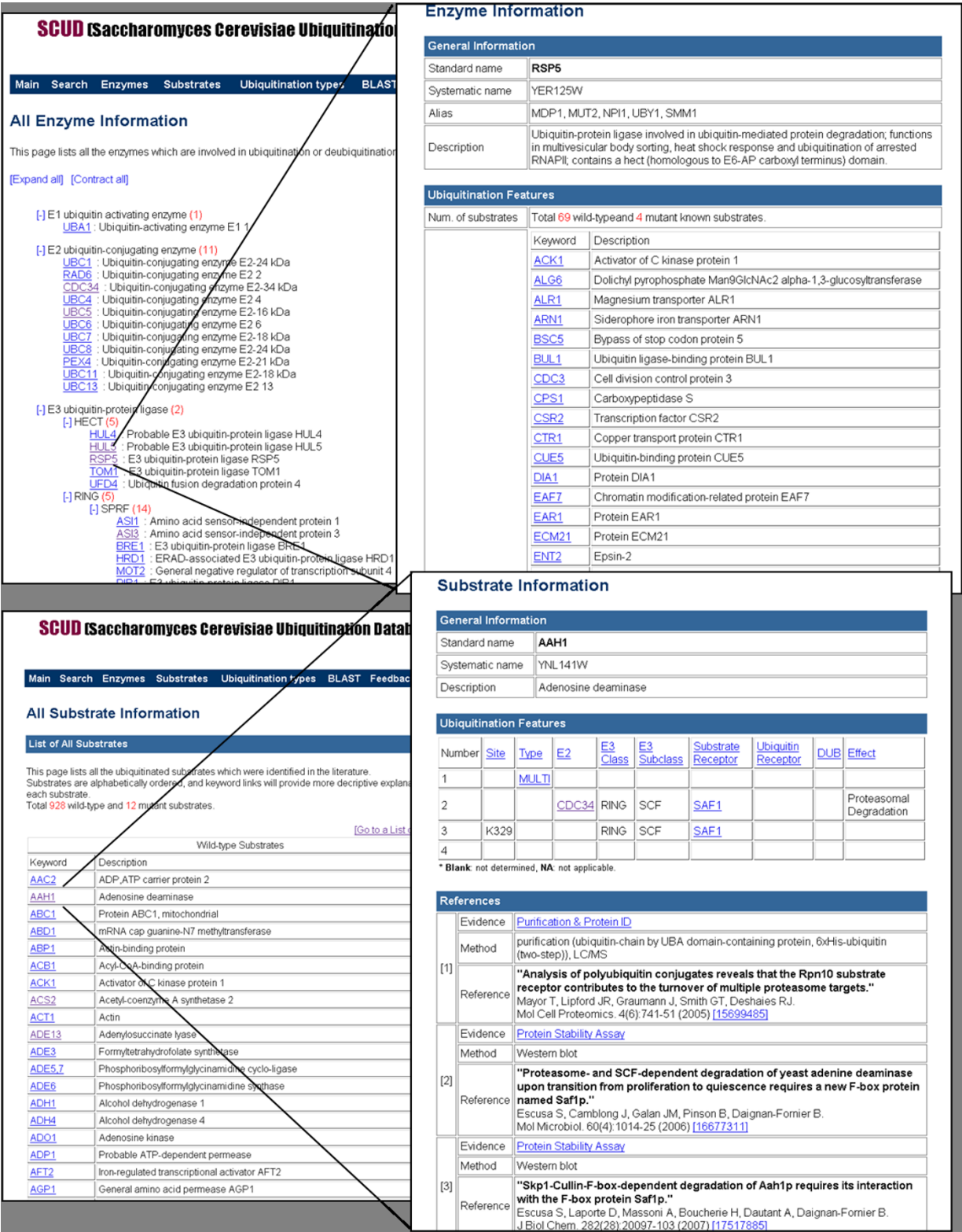
The numbers of E3 enzymes for each E3 class and subclass. Total 42 E3 enzymes are sorted into different E3 classes and subclasses.

detailed methods and related reference information (Figure 3). It will help users to examine the ubiquitinated proteins with their own filtering criterions.

Finally, the database enables users to discriminate the substrates by their type of ubiquitination such as mono- and Lys48-linked ubiquitination (Figure 4). We also provide BLAST search service which runs on all the protein sequences including enzymes and substrates in our database. Moreover, researchers in ubiquitination field can contribute to the database by submitting new proteins including enzymes and substrates, by finding errors to be corrected, and by giving general opinions and suggestions about the database itself (feedback session). The database will be regularly updated with additional data and corrections.

Conclusion

SCUD is a web-based database aiming to represent a comprehensive ubiquitination system for yeast. We identified all known ubiquitination enzymes including E1, E2, E3, and deubiquitination enzymes as well as substrates by literature search. Users can examine various ubiquitination features such as ubiquitination site, type, and effect, and obtain enzyme-specific or ubiquitination type-specific information by hyperlinks. This database is easily expandable with further experimental data. Built-in BLAST search is useful when users want to test putative protein sequences that may belong to the ubiquitination system of yeast or other species. The feedback interface will gather a variety of ubiquitination information as well as opinions and suggestions about the database itself. We expect that it will be useful for the researches on other higher organisms.



SCUD (Saccharomyces Cerevisiae Ubiquitination Database)

[Main](#) [Search](#) [Enzymes](#) [Substrates](#) [Ubiquitination types](#) [BLAST](#) [Feedback](#) [HELP](#)

All Ubiquitination Type

List of all ubiquitination types

This page lists all the ubiquitination types annotated to the substrates in the database.

Total 6 ubiquitination types.

Type	Description
MONO	Modification of a lysine residue with a single ubiquitin moiety.
PLURI	Multisite mono-ubiquitination (modification of several lysine residue with a single ubiquitin moiety).
MULTI	Several ubiquitins form a ubiquitin chain, but its exact linkage such as Lys48-linked is unknown.
MULTI(K48-linked)	Ubiquitins form a ubiquitin chain by linkage through 48th lysines.
MULTI(K63-linked)	Ubiquitins form a ubiquitin chain by linkage through 63rd lysines.
AUTO	Substrate receptor type is same as its substrate.

[Main](#) [Search](#) [Enzymes](#) [Substrates](#) [Ubiquitination types](#) [BLAST](#) [Feedback](#) [HELP](#)

Ubiquitination Type Information

General Information

Ubiquitination Type	MONO
Description	Modification of a lysine residue with a single ubiquitin moiety.

Substrates with MONO

Keyword	Number	E2	E3 Class	E3 Subclass	Substrate Receptor	Ubiquitin Receptor	DUB	Effect
GIS4	1		RING	SCF	GRR1			
HTB2	1	RAD6	RING	SPRF	BRE1		UBP8	
MAL61	1	UBC1 , UBC4 , UBC5	HECT	NA	RSP5			
PGI1	1		HECT	NA	RSP5			
POL30	2	RAD6	RING	SPRF	RAD18			Error-prone Pathway of TLS (Translesion Synthesis) DNA repair
RPB7	1		HECT	NA	RSP5			
RVS167	2		HECT	NA	RSP5			
STE2	1							
STE2	2							
STE2	3	UBC1 , UBC4 , UBC5	HECT	NA	RSP5			Vacuolar Degradation
VPS30	1		HECT	NA	RSP5			

Figure 4
Ubiquitination types. A list of all identified ubiquitination types and ubiquitination type-specific information.

Availability and requirements

SCUD is available at <http://scud.kaist.ac.kr/>

License: The database is freely available.

Authors' contributions

WCL carried out a literature search, participated in database construction, and drafted the manuscript. ML participated in programming for database construction and preparation of the manuscript. JWJ designed an overall database scheme. KPK and DK supervised the work and collaborated in writing the manuscript. All authors have read and approved the manuscript.

Additional material

Additional file 1

Ubiquitinated wild-type proteins in SCUD. Total 928 wild-type substrates with various ubiquitination features such as ubiquitination site, type, enzymes, receptors, and effect. Experimental evidences and detailed methods are also included.

Click here for file

[<http://www.biomedcentral.com/content/supplementary/1471-2164-9-440-S1.txt>]

Acknowledgements

This work is supported by Korea Institute of Science and Technology Information (KISTI) Supercomputing Center. ML was supported by Kim Bo Jeong Basic Science Scholarship of Korea Advanced Institute of Science and Technology (KAIST). KPK was supported by a grant of the Korean Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (A060769).

References

- Chernorudskiy AL, Garcia A, Eremin EV, Shorina AS, Kondratieva EV, Gainullin MR: **UbiProt: a database of ubiquitylated proteins.** *BMC Bioinformatics* 2007, **8**: Article No: 126
- Huibregtse JM, Scheffner M, Beaudenon S, Howley PM: **A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase.** *Proc Natl Acad Sci USA* 1995, **92**(11):5249.
- Joazeiro CA, Weissman AM: **RING finger proteins: mediators of ubiquitin ligase activity.** *Cell* 2000, **102**(5):549-552.
- Hatakeyama S, Nakayama KI: **U-box proteins as a new family of ubiquitin ligases.** *Biochemical and biophysical research communications* 2003, **302**(4):635-645.
- Koegl M, Hoppe T, Schlenker S, Ulrich HD, Mayer TU, Jentsch S: **A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly.** *Cell* 1999, **96**(5):635-644.
- Hitchcock AL, Auld K, Gygi SP, Silver PA: **A subset of membrane-associated proteins is ubiquitinated in response to mutations in the endoplasmic reticulum degradation machinery.** *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(22):12735-12740.
- Peng J, Schwartz D, Elias JE, Thoreen CC, Cheng D, Marsischky G, Roelofs J, Finley D, Gygi SP: **A proteomics approach to understanding protein ubiquitination.** *Nature Biotechnology* 2003, **21**(8):921-926.
- Kus B, Gajadhar A, Stanger K, Cho R, Sun W, Rouleau N, Lee T, Chan D, Wolting C, Edwards A, et al.: **A high throughput screen to identify substrates for the ubiquitin ligase Rsp5.** *Journal of Biological Chemistry* 2005, **280**(33):29470-29478.
- Mayor T, Lipford JR, Graumann J, Smith GT, Deshaies RJ: **Analysis of polyubiquitin conjugates reveals that the Rpn10 substrate receptor contributes to the turnover of multiple proteasome targets.** *Molecular & Cellular Proteomics* 2005, **4**(6):741-751.
- Kostova Z, Wolf DH: **For whom the bell tolls: protein quality control of the endoplasmic reticulum and the ubiquitin-proteasome connection.** *The EMBO journal* 2003, **22**(10):2309-2317.
- Johnson ES, Ma PC, Ota IM, Varshavsky A: **A proteolytic pathway that recognizes ubiquitin as a degradation signal.** *The Journal of biological chemistry* 1995, **270**(29):17442-17456.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

